

基于 SOA 的农作物种质资源调查信息系统研究

司海平, 方 洵, 唐 鹏, 曹永生
(中国农业科学院作物科学研究所, 北京 100081)

摘要: 在研究面向服务体系结构(SOA)基础上,采用价值链分析方法识别了农作物种质资源调查服务域,提出基于 SOA 的农作物种质资源调查信息系统体系结构参考模型,详述了参考模型的层次结构,定义了各个服务模块功能。该参考模型为系统进一步编码和实现奠定了基础。

关键词: 面向服务体系架构; 农作物; 种质资源; 调查

Research of Crop Germplasm Resources Investigation Information System Based on SOA

SIHaiping FANG Wej TANG Peng CAO Yong-sheng
(Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract Crop germplasm resources investigation is the basis of germplasm resources management. Based on the analysis of the features of service-oriented architecture(SOA), the services domain in crop germplasm resources investigation were identified by the value chain analysis and the reference model of SOA-based crop germplasm resources investigation information system (CGRIS) was proposed. The hierarchical structure of the reference model and the definitions of each service module functions were described. The reference model laid the foundation for system further coding and implementation.

Key words SOA; Crop Germplasm resources; Investigation

农作物种质资源调查是农作物种质资源工作的基础,对于农作物种质资源的保护和利用具有重要意义。通过农作物种质资源调查可以收集农作物种质材料,获取资源的分布和数量等重要信息。20世纪50年代以来,我国在农作物种质资源调查方面做了大量工作,先后组织了多次农作物种质资源调查,收集到大量的种质资源,在资源调查和信息共享方面取得了重大进展和显著成效^[1]。但农作物种质资源调查也存在一些突出问题,如缺乏统一的调查数据标准和规范、调查信息化程度不高、未有效整合不同项目的调查数据、缺乏全国性的农作物种质资源调查信息网络服务系统等。

针对以上问题,在分析农作物种质资源调查和面向服务(service-oriented architecture, SOA)特点的基础上,本文设计了基于 SOA 的农作物种质资源调

查信息系统(crop germplasm resources investigation information system, CGRIS),以达到农作物种质资源调查数据标准统一、信息整合和平台通用的目标,规范农作物种质资源调查工作。

1 SOA

SOA 是一个组件模型,它将应用程序的不同功能单元(称为服务)通过定义良好的接口和契约联系起来,使得构建在基于 SOA 系统中的服务可以以一种统一和通用的方式进行交互^[2]。W3C(world wide web consortium)将 SOA 定义为:“一种应用程序体系结构,在这种体系结构中,所有的功能都定义为独立的服务,这些服务带有定义明确的可调用接口,采用已定义好的顺序调用这些服务来形成业务

收稿日期: 2009-11-19 修回日期: 2010-06-11

基金项目: 国家科技基础工作专项(2006FY110700)

作者简介: 司海平, 博士研究生, 研究方向为作物信息技术、计算机网络。E-mail: shaping@yahoo.com.cn

通讯作者: 曹永生, 研究员, 主要从事种质资源信息系统、图像识别、GIS 领域研究。E-mail: caoy@mail.caas.net.cn

流程”^[3]。从本质上讲, SOA 是一种粗粒度、松耦合服务架构, 服务之间通过简单、精确定义接口进行通讯, 不涉及底层编程接口和通讯模型, 这样服务可以以一种统一和通用的方式进行交互。

1.1 SOA 模型

典型的 SOA 模型如图 1 所示, 包括服务提供者、服务请求者和服务注册中心 3 种角色^[3]。服务都需要在服务注册中心按照服务注册合约进行登记, 服务注册合约中除了服务合约外, 还包括授权、使用条件以及服务提供者等信息。

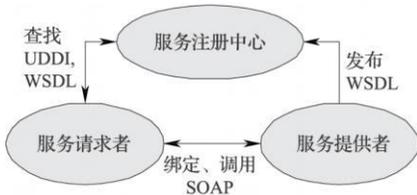


图 1 SOA 模型

Fig 1 SOA model

服务提供者 (service provider) 将可用的过程或信息封装成服务, 用独立于平台和语言的方式定义对内和对外的接口, 再把接口的描述信息发布到服务注册中心以供服务请求者使用。

服务注册中心 (service registry) 接受服务提供者注册发布的服务, 并对已经注册的服务进行分类, 同时为服务请求者 (service consumer) 提供服务描述信息的查询; 通过验证后, 向服务请求者提供服务地址等信息。服务提供者接到请求后按合约提供服务。

Web service 是 SOA 的方式之一^[4], 它提供的基于可扩展标记语言 (extensible markup language XML) 标准化接口能较好支持 SOA 应用模型的实现。

Web 服务描述语言 (web service definition language WSDL)^[5] 用来描述服务。统一描述、发现和集成协议 (universal description discovery and integration UDDI)^[6] 用来注册和查找服务。简单对象访问协议 (simple object access protocol SOAP)^[7] 作为传输层, 是 Web 服务的默认机制, 用来在服务请求者和提供者之间传送信息。UDDI 注册表查找服务, 取得服务的描述, 然后通过简单 SOAP 来调用服务。

1.2 SOA 设计过程

SOA 的实现采用自顶向下的设计方法, 通过业务流程执行语言 (business process execution language BPEL) 灵活且独立地编写新的业务流程, 达到随需应变的目的。在设计基于 SOA 系统时, 一切都是从服务的角度出发, 首先考虑服务需求, 进行系统服务组合、装配。整个软件过程划分为服务描述、

服务发布、服务发现、服务绑定、业务逻辑设计和服务执行过程^[8-9]。软件的功能模块体现为服务组件、接口。软件的设计过程如图 2 所示。

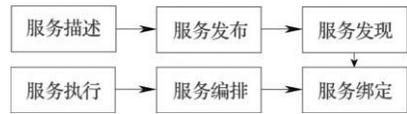


图 2 SOA 架构设计过程

Fig 2 SOA architecture design process

2 CGR IIS 总体架构设计

CGR IIS 采用基于 SOA 设计方法, 围绕服务构建系统, 使系统更适应调查项目变化, 满足不同种质资源调查项目需要。

2.1 CGR IIS 价值链分析

价值链分析法利用系统的方法来观察各项活动及其相互关系, 从而找出能够增加系统价值的活动, 识别系统的主要任务和目标, 同时也可以识别系统所需的各种服务^[10-11]。价值代表要满足用户的需求。价值活动代表可以创造出对用户有价值的系统构成要素, 不同系统的价值活动划分与构成不同, 得出的价值链也不同。价值活动可以分为基本活动和辅助活动两大类。基本活动指与系统要满足的需求或服务直接相关的活动。辅助活动为基本活动的执行提供辅助支持。辅助活动的每一项都可与具体的基本活动联系起来并支持整个价值链。图 3 是 CGR IIS 价值链分析结果。

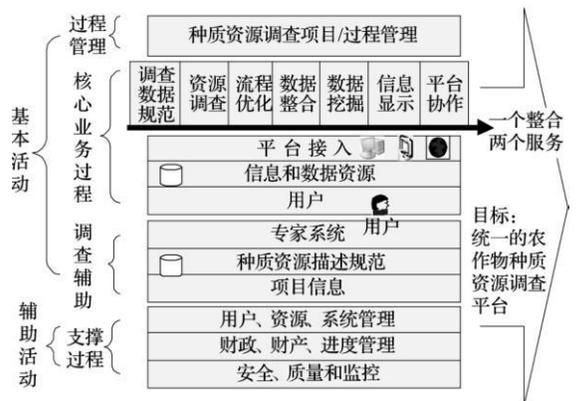


图 3 CGR IIS 价值链

Fig 3 CGR IIS value chain

在 CGR IIS 的价值链分析中, 基本活动在这里是指完成资源调查数据录入和项目间数据整合所必需的活动, 这些活动与系统功能的实现直接相关。辅助活动是维持项目运行所需的活动, 但该类活动不能增加项目或服务的价值。如图 3 所示, CGR IIS 价值链的基本活动由 7 个可以增加价值的活动组成: 调查数据规范, 资源调查, 流程优化, 数据整合, 数据挖掘, 信

息显示,平台协作。不同活动通过链接、组合达到资源调查、信息收集、数据整合和数据综合处理等目的。通过价值链分析, CGR IIS 要达到一个整合两个服务的目的: 整合不同调查项目间的数据, 使调查信息得到更充分的利用; 为调查人员提供详细的网络系统支撑服务, 优化资源调查流程, 使种质资源调查更标准化和现代化; 为上级决策部门和科学研究提供信息服务, 提供多层次、多角度的数据分析及展现, 最终达到建立全国统一农作物种质资源调查信息系统的目标。

2.2 种质资源调查过程模型

农作物种质资源调查数据录入和不同调查项目间数据整合是 CGR IIS 的主要功能, 图 4 给出了用户使用 CGR IIS 进行调查数据录入、项目间调查和数据整合的业务过程模型, 在该模型中主要包含两种用户: 资源调查用户和项目用户。资源调查用户首先要取得 CGR IIS 用户验证, 然后通过项目选择、数据录入、数据规范检查、数据修改、数据存储等过程完成调查资源数据的录入过程。项目用户包括资源调查项目负责人或与资源调查项目相关的管理人员。项目用户通过项目管理功能来定制、修改、完善项目, 管理项目进展、存储技术资料等工作。统计整合数据信息后以文字和图形图像的方式呈现给用户。

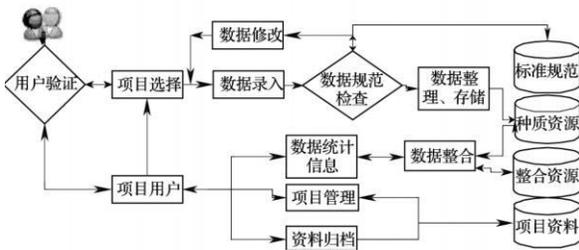


图 4 种质资源调查和数据整合过程模型

Fig 4 Gemplasm resources investigation and data integration process model

2.3 基于 SOA 的 CGR IIS 解决方案

农作物种质资源调查信息系统 (CGR IIS) 设计基于 SOA 模式, 服务通过 Web Services 技术实现。系统各模块的功能都定义为独立的服务, 调查业务流程通过工作层单个服务或者多个服务的组合完成, 从而为表示层提供了一个整体的业务逻辑视图。提高重用性是 CGR IIS 架构设计的一个重要目标, CGR IIS 的各种服务以应用程序编程接口 (application programming interface API) 形式给出, 以一种松散耦合的方式集成, 来达到重用目的。应用程序编程接口通过 WSDL 描述, 并使用 SOAP 传输方式作为与外界交互的桥梁。系统的整体架构和服务部署如图 5 所示。

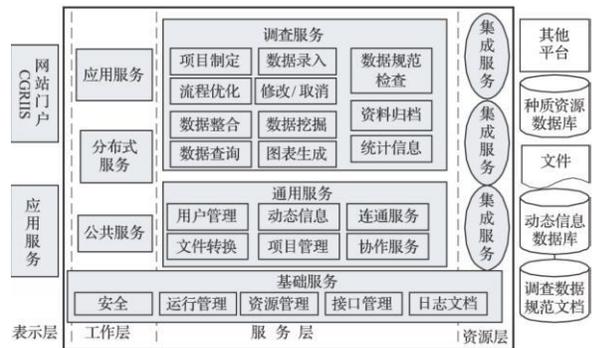


图 5 CGR IIS 服务架构

Fig 5 CGR IIS Services architecture

2.3.1 表示层 表示层包括应用服务和 CGR IIS 网站门户两部分, CGR IIS 网站门户提供统一身份认证、登录、界面定制、应用访问、数据展现等功能。应用服务包括接入本系统的其他应用程序和 web 站点接入。系统用户包括调查人员、管理人员、其他 Web 站点、普通浏览用户、系统维护人员和相关专家等。

2.3.2 工作层 工作层包括分布式服务、应用服务和公共服务。应用服务用来管理应用逻辑如调查过程等。分布式服务管理任务、会话状态、操作授权和认证等工作。公共服务提供一般性业务功能, 如项目信息, 动态信息查询等功能。

2.3.3 服务层 服务层是 CGR IIS 设计工作的重点, 提供种质资源调查和其他工作所必需的服务。这些服务分为调查服务、通用服务和基础服务三类。

(1) 调查服务。提供与具体资源调查相关的功能服务, 具体包括以下几个方面。

项目定制 支持项目人员定制不同农作物种质资源调查项目, 项目采用统一调查数据标准和规范, 不同项目对调查的数据项有不同选择。有 3 种不同的调查项目, 即标准项目、定制项目和条件项目。

标准项目是一些重大的种质资源调查项目, 该项目的调查数据项默认必选。定制项目系为一些特殊调查活动所定制的项目, 调查数据项是可以选择的。条件项目为某一类特定调查项目, 在该调查项目中某些数据项是必选的。

数据录入 调查人员按照统一的农作物种质资源调查数据标准输入项目调查数据。该服务提供多种录入方式, 可以采取批处理的方式通过接口一次性输入多个调查数据, 也可以通过移动终端 (移动电话、手机、PDA、GPS 等) 采集设备录入。数据包括字符、数字、图片、视频和 GIS 等资料。

修改 删除 用户对自己输入的调查数据进行修改或删除操作。

数据规范检查 检查资源调查数据是否符合统一的农作物种质资源调查标准规范要求。在数据采集过程中要严格按照农作物种质资源基本描述规范要求填写数据。如种质资源的全国统一编号规定由8位字符串组成,如“XG001035”,编号具有全国唯一性^[12]。统一的调查数据标准有利于整合全国的农作物种质资源,规范农作物种质资源的收集工作。

数据整合 整合不同项目间的种质资源调查数据。整合后的信息另外存储来保持项目调查数据独立性,使调查信息得到更充分利用。

流程优化 提供调查项目相关的信息,通过优化调查地区优先次序、调查路线和取样策略等方法优化资源调查流程。

数据挖掘 对调查得到的种质资源数据进行进一步的挖掘和利用。

资料归档 归档调查过程中的调查方法、收集技术总结、遇到的特殊情况、考察中的经验和教训、种质鉴定结果等文字资料。

数据查询 根据不同条件(如属性等)查询项目调查数据详细资料。

统计信息 分析资源的分布和数量关系,对各种资源调查数据按照特定条件进行统计和显示。该服务支持多种方式显示的统计信息,如饼形图、柱状图、条形图、折线图等。

(2)通用服务。提供资源调查项目相关的功能服务,包括用户管理、项目管理、连通服务、文件转换、动态信息、协作服务。

用户管理 对用户进行认证和授权,实现用户与角色、角色与功能的对应管理,实现分级授权、功能代理、权限委托等多种权限的分发管理和调用功能。

动态信息 提供项目、业务、政策、法规等相关的动态新闻和信息。可以通过信息订阅的方式,通过E-Mail或简易信息聚合(really simple syndication RSS)等方式订阅门户网站的相关内容。

连通服务 维持用户连接的持续性。

项目管理 包括对调查项目具体调查数据信息进行管理,如搜索、查询和修改特定资源调查数据等操作,也包括对调查人员、项目资金的管理等工作。

文件转换 提供多种格式(如文本文档)的数据信息输出方式。

(3)基础服务。提供维持CGRIS系统运行所必须的基础性服务,包括接口管理、资源管理、运行管理、数据安全存储、备份和加密,操作日志文档和其他帮助信息。

2.3.4 资源层 CGRIS资源层包括种质资源调查数据库、调查数据规范文档、动态信息数据库等。其他平台通过服务层中协作服务来实现与本系统的相互操作。集成服务用来连接不同的数据库,为服务提供不同的数据支持,包括种质资源数据库、资源调查数据规范文档、动态信息数据库等。

2.4 CGRIS开发方案

系统采取J2EE开发平台。通过无状态EJB(Enterprise Java bean)组件,系统把各种服务和业务流程公开为Web服务,Web服务用户也可以通过EJB服务接口访问无状态会话应用程序(Bean)调用服务。在J2EE平台上实现Web service可以充分利用现有的业务逻辑和流程,发挥Web服务松散耦合、即时集成的优势,又可以充分利用J2EE标准中EJB组件的可复用性、多层分布企业解决框架的优势。

3 讨论

计算机和网络技术的快速发展奠定了农作物种质资源调查信息化的基础,而面向服务架构(SOA)通过屏蔽后台技术的复杂性为软件体系结构设计提供了一种灵活方案。本文采用价值链分析方法识别了农作物种质资源调查服务域,设计了基于SOA的CGRIS体系结构,详述了各服务的功能和交互关系,并给出了系统开发方案。该架构把与种质资源调查相关的服务作为基本元素,服务可在不同调查流程中被重复使用,使系统具有开放、灵活、可扩展、可重用的特点。系统的建立将极大地促进种质资源调查工作的标准化和信息化。

农作物种质资源调查数据的特殊性要求系统必须确保数据存储和传输的安全性,这也是系统需要进一步解决的问题。

参考文献

- [1] 卢新雄,曹永生.作物种质资源保存现状与展望[J].中国农业科技导报,2001,3(3):43-47
- [2] 郝轩,史维峰.基于SOA的Web服务动态组合研究与实现[J].计算机工程与设计,2008,29(8):3929-3932
- [3] Thomas Erl Service-Oriented Architecture(SOA): Concepts Technology, and Design [M]. The PrenticeHall 2005
- [4] Eric Newcomer Greg Lomow. Understanding SOA with Web Service[M]. Independent Technology Guides David Chapell Series Editor 2005
- [5] Roberto Chimicj Jean- Jacques Moreau, Arthur Rynan Web services description language(WSDL) version 2.0 part 1 core language[EB/OL]. (2005-8-3)[2009-10-1]. <http://www.w3.org/TR/2005/WD-wsdl20-20050803>
- [6] Luc Clement Andrew Hatley Klaus von Riegen, et al UDDI Version 3.0.2[EB/OL]. (2004-10-19)[2009-10-12]. http://www.uddi.org/pubs/uddi_v3.htm

- [7] Martin Gudgin M arc Hadley, Jean- Jacques Moreau et al SOAP Version 1. 2 [EB/OL]. (2001- 7- 9) [2009- 10- 6]. <http://www.w3.org/TR/2001WD-soap12-20010709>
- [8] Daniel Austin W W Grainger Abbie Barbic et al Web services architecture requirements [EB/OL]. (2002- 4- 29) [2009- 10- 10]. <http://www.w3.org/TR/2002WD-wsa-reqs-20020429>
- [9] Fehia Rahji Hairong Yu, Feras T Dabous et al A service-oriented architecture for financial business processes Information Systems and E- Business Management 2007, 5(2): 185- 200
- [10] 齐二石, 王慧明. 基于价值链理论的业务流程再造研究 [J]. 工业工程, 2005(1): 57- 60
- [11] 王田苗, 胡耀光. 基于价值链的企业流程再造与信息集成 [M]. 北京: 清华大学出版社, 2002
- [12] 刘旭, 曹永生, 张宗文. 农作物种质资源基本描述规范 [M]. 北京: 机械工业出版社, 2008 3 4- 6
- [13] A Arsanjani L- J Zhang M Ellis et al Design an SOA Solution Using a Reference Architecture [EB/OL]. [2007- 03]: <http://www.ibm.com/developerworks/library/ar-archtemp>
- [14] 魏东, 陈晓江, 房鼎盛. 基于 SOA 体系结构的软件开发方法研究 [J]. 微电子学与计算机, 2005, 22(6): 73-76
- [15] 杨涛, 刘锦德. Web Services技术综述——一种面向服务的分布式计算模式 [J]. 计算机应用, 2004, 24(8): 1-4
- [16] 麻志毅, 陈泓斌. 一种面向服务的体系结构参考模型 [J]. 计算机学报, 2006 29(7): 1011- 1019
- [17] 曹永生, 陈育, 孔繁胜. 中国作物种质资源信息共享网络的建立 [J]. 资源科学, 2001, 23(1): 46- 48

欢迎
订阅

《作物学报》是中国科学技术协会主管、中国

作物学会和中国农业科学院作物科学研究所共同主办、科学出版社出版的有关作物科学方面的学术期刊。主要

刊载农作物遗传育种、耕作栽培、生理生化、种质资源以及与作物生产有关的生物技术、生物数学等学科具基础理论或实践应用性的原始研究论文、专题评述和研究简报等。办刊宗旨是报道本领域最新研究动态和成果, 为繁荣我国作物科学研究、促进国内外学术交流、加速中国农业现代化建设服务。读者对象是从事作物科学研究的科技工作者、大专院校师生和具有同等水平的专业人士。

《作物学报》为月刊, 2011年定价 50元/册, 全年 600元。可通过全国各地邮局订阅, 刊号: ISSN 0496-3490 CN 11-1809/5 邮发代号: 82-336 也可向编辑部直接订购。

地址: (100081)北京市海淀区中关村南大街 12号 中国农科院作物所《作物学报》编辑部

电话: 010-82108548 传真: 010-82105793

E-mail: xbzw@chinajournal.net.cn

网址: <http://www.chinacrops.org/zwxzb/>

《云南植物研究》(2011年拟更名为“植物分类

与资源学报”)是由中国科学院主管、中国科学院昆明植物研究所承办的全国性自然科学学术期刊, 中国自然科学核心期刊、中国生物学类科技核心期刊。本刊荣获中科院优秀期刊二等奖(1996)及一等奖(2000)、第二届全国优秀期刊三等奖(1997)及云南省优秀科技期刊一等奖(1997)、提名奖(2007)、优秀奖(2009)等, 并作为中国科学院首批向美国 SC I推荐的刊物之一, 2002年入选国家“双效期刊”。本刊所发表的论文在国内生物、农林、医药、轻工等二次文献刊物都有摘报; 国外 CA(美国化学文摘)、BA(美国生物学文摘)等从 1980年起就连续摘报; 生物科学的当代进展(CABS)、科学引文索引(SCI)的 C I部分以及俄罗斯文摘杂志(P)和国际农业科技情报系统(Agris)等都有摘报; 乌利希国际期刊指南(UPI)从 20世纪 80年代就刊载本刊出版事宜。本刊主要报道植物学各分支学科具有创造性或较高学术水平的研究论文和简报; 植物学领域的新发现及重大应用价值的新成果; 有关植物学资源开发利用和保护的创新性研究成果; 植物学研究的新技术、新方法; 反映本学科重要领域的国内外植物科学研究的最新进展的评述, 中英文稿件均受欢迎。

《云南植物研究》为双月刊, 双月 25日出版, 2011年每期 25元, 邮发代号: 64-11 若在邮局漏订的读者可直接与编辑部联系订阅。

联系地址: (650204)云南昆明市北郊黑龙潭 中国科学院昆明植物研究所

E-mail: bjnj@mail.kh.ac.cn <http://journal.kh.ac.cn>

Tel& Fax 0871-5223032

《遗传学报》和《遗传》杂志是中国遗传学会和中国科学院遗传与发育生物学研究所主办、科学出版社出版的学术期刊, 中国精品科技期刊, 已被美国化学文摘(CA)、生物学数据库(BIOSIS)、生物学文摘(BA)、医学索引(Medical Index)、俄罗斯文摘杂志(AJ)以及NCBI/CABI等 20多种国内外重要检索系统与数据库收录。刊登内容涉及遗传学、发育生物学、基因组学、细胞生物学以及分子进化等。读者对象为基础医学、农林牧渔、生命科学领域的科研与教学人员、研发人员、研究生、大学生、中学生物学教师等。

《遗传学报》(ISSN 1673-8527 CN11-5450/R)为月刊, 全年 12期, 国内邮发代号 2-819, 国外发行代号: M63, 2011年定价 50元, 全年 600元。 <http://www.jgenetgenomics.org>

《遗传》(ISSN 0253-9772 CN11-1913/R)为月刊, 全年 12期, 国内邮发代号 2-810 国外发行代号: M62, 2011年定价 50元, 全年 600元。 <http://www.Chinagene.cn>

联系地址: (100101)北京市朝阳区北辰西路 1号院: 中国科学院遗传与发育生物学研究所

电话: 010-64807669

传真: 010-64807786

《植物资源与环境学报》 中国科技核心期刊

中国科学引文数据库核心期刊 季刊, 单价 15元, 邮发代号: 28-213, 国内统一连续出版物号: CN 32-1339/S

系江苏省·中国科学院植物研究所、江苏省植物学会及中国环境科学学会植物园保护分会联合主办的学术刊物, 围绕植物资源与环境两个中心命题, 报道我国植物资源的考察、开发利用和植物物种多样性保护, 自然保护区与植物园的建设和管理, 植物在保护和美化环境中的作用, 环境对植物的影响以及与植物资源和植物环境有关学科领域的原始研究论文、研究简报和综述等。

本刊于 1992年创刊, 全国各地邮局发行, 若错过征订时间或需补齐 1992年至 2010年各期者, 请直接与编辑部联系邮购。邮购价: 1992年至 1993年每年 8元; 1994年至 2000年每年 16元; 2001年至 2005年每年 24元; 2006年至 2008年每年 40元; 2009年至 2011年每年 60元(均含邮资, 如需挂号另付挂号邮资费 3元)。

编辑部地址: (210014)南京中山门外·中国科学院植物研究所内

电话: 025-84347016 Fax 025-84432074

E-mail: zwzy@mail.cnbg.net.cn 或 nbgx@jonline.cn